

ОСОБЛИВОСТІ ОБРОБКИ ФАЗОВОЇ ІНФОРМАЦІЇ ГОЛОСОВОГО СИГНАЛУ КОРИСТУВАЧА СИСТЕМИ ГОЛОСОВОЇ АУТЕНТИФІКАЦІЇ

Пастушенко М.С., Красноженюк Я.О., Заїка М.В.

Кафедра «Інфокомунікаційної інженерії
ім. В.В. Поповського», ХНУРЕ, Україна.

E-mail: mykola.pastushenko@nure.ua
yana.krasnozheniuk@nure.ua
maksym.zaika@nure.ua

Abstract

A relevant scientific problem of improving the quality of voice authentication systems, which have recently been subjected to intensive scientific research, is considered. The paper analyzes the influence of the phase data of the user's voice signal, as well as the influence of the sampling frequency used on the quality of generated attributes of an authentication system template. A research methodology has been developed, which allows to assess the potential impact of the sampling rate on the quality of the formation of user template elements. The results obtained can be useful in the processing of speech signals in other areas, for example, forensics, separation of speakers, the creation of voice control systems for computers, etc.

Загальна постановка задачі

Вплив інформаційної сфери на розвиток сучасного суспільства безперервно зростає. У зв'язку з цим забезпечення інформаційної безпеки стає одним з пріоритетів національної політики будь-якої держави.

Зміст проблем, що об'єднуються поняттям «інформаційна безпека», останній часом визначається, перш за все, швидким поширенням нових інформаційно-комунікаційних технологій. Ці технології, що призвели до проникнення комп'ютеризації в усі сфери людської діяльності, збільшують залежність від інформаційних систем і послуг і створюють нові види загроз для інтересів окремої особи, підприємств і організацій, суспільства в цілому.

Уразливість приватних осіб, організацій і держави по відношенню до загроз інформаційної безпеки особливо зростає при використанні телекомунікаційних мереж — як загального користування, так і корпоративних. Цьому сприяє також поширена тенденція до розподіленої обробки даних, яка пов'язана з використанням дистанційного режиму і телекомунікаційних технологій (зокрема, розширюється сфера діяльності співробітників і залучених осіб поза межами відповідної організації).

Все більших масштабів набувають кримінальні напрямки комп'ютерної діяльності. Згідно з уніфікацією Комітету міністрів Європейської Ради, до цих напрямків можна віднести комп'ютерне шахрайство, несанкціонований доступ до інформації, підробку комп'ютерної інформації, несанкціоноване перехоплення даних і інші види злочинних дій [1]. У зв'язку з цим найважливішим завданням стає створення і застосування нових ефективних методів і засобів захисту інформації.

Розвиток нових методів і засобів забезпечення інформаційної безпеки покликаний, перш за все, запобігти загрозам доступу до інформаційних ресурсів сторонніх осіб, які не мають доступу. Для вирішення цього завдання необхідна наявність ідентифікаторів і створення процедур ідентифікації для всіх користувачів [2].

Сучасні ідентифікація і аутентифікація включають в себе різні системи і способи біометричної ідентифікації особи. Розвиток систем ідентифікації особи, заснованих на біометричних вимірюваннях, пов'язаний з цілим комплексом переваг: такі системи більш надійні, оскільки біометричні показники складніше підробити; сучасна мікропроцесорна техніка робить біометричні методи (невід'ємні біометричні ідентифікатори) більш зручними в порівнянні зі звичайними методами ідентифікації

(електронними ідентифікаторами); нарешті, вони значно простіше піддаються автоматизації вимірювань.

Однією з найбільш поширених біометричних характеристик людини є її голос, що володіє набором індивідуальних особливостей, які відносно легко піддаються вимірюванню (наприклад, частотний спектр голосового сигналу). До переваг голосової ідентифікації відносяться також зручність застосування і використання, досить невисока вартість пристроїв, застосовуваних для ідентифікації (наприклад, мікрофонів).

Можливості ідентифікації особи за голосовими даними охоплюють досить широкий спектр завдань, що виділяє їх серед інших біометричних систем. Перш за все, голосова ідентифікація досить давно і широко використовується в різних системах розмежування доступу до фізичних об'єктів і інформаційних ресурсів [3, 4].

Існує багато інструментів дослідження мовних сигналів, заснованих на математичній обробці, серед яких основним є спектральний аналіз. Використання часової залежності і спектрального аналізу дає додаткову інформацію і називається спектрально-часовим аналізом голосових сигналів.

При цьому в даний час все частіше починають враховувати фазові характеристики голосових сигналів [5, 6]. До того ж виникає ряд нових актуальних наукових завдань, які раніше не розглядалися, наприклад: процедури попередньої обробки фазових даних і їх вплив на якість характеристик систем голосової аутентифікації; вплив використовуваної частоти дискретизації вхідного голосового сигналу на якість формованих фазових даних; використання фазових даних для проведення коригування матеріалів реєстрації голосового сигналу. У даній роботі аналізується вплив використовуваної частоти дискретизації голосового сигналу на якість формування фазових даних і характеристики процедур попередньої обробки матеріалів реєстрації.

Методика і результати досліджень

При формуванні ознак шаблону користувача необхідно виконати попередню обробку матеріалів реєстрації, які включають процедури вирівнювання аналізованих сигналів за амплітудою і часом. Вирівнювання сигналів за амплітудою вирішується досить просто, наприклад, шляхом нормування матеріалів реєстрації на максимальне значення. Більш складно вирівняти аналізовані сигнали за часом. У найпростішому випадку в якості аналізованого сигналу вибирається однакове значення відліків аналізованого сигналу. Цей метод має істотний недолік, він призводить до зміни змісту аналізованих даних, які впливають на формування шаблону. Більш складні процедури часового вирівнювання передбачають передискретизацію матеріалів реєстрації.

В роботі розглядається інший підхід часового вирівнювання матеріалів реєстрації, який передбачає реєстрацію голосового сигналу на високій частоті дискретизації. Наприклад, в даному дослідженні частота дискретизації сигналу становила 64 кГц. Аналізу піддавався мовний сигнал користувача цифр від 0 до 9. Відношення сигнал/шум аналізованої послідовності становило понад 25 дБ. При цьому оцінювалися ознаки шаблону користувача, які знайшли широке застосування на практиці: частота основного тону, формантна інформація, спектральні і кепстральні коефіцієнти. Як математичний апарат аналізу голосового сигналу використовувався спектральний аналіз. При цьому основна увага приділялася аналізу діапазону спектра до 8 кГц. Останнє обумовлено наявністю характерних ознак користувача на цій ділянці спектра (в діапазоні від 0,1 кГц до 8 кГц).

Методика проведених досліджень передбачала формування фазових даних, які, на жаль, у голосового сигналу не реєструються. У сучасних системах цифрової обробки даних формування фазових даних базується на використанні перетворення Гільберта.

Далі необхідно виконати коригування фазових даних, оскільки функція арктангенс формує некоректний кут. Тут же зауважимо унікальну особливість фазових даних голосового сигналу - це апріорно відомі форма фазового сигналу і межі його зміни. Останнє дає підставу використовувати фазові дані і для коригування амплітудної та частотної інформації голосового сигналу, що реєструється. Така попередня обробка матеріалів реєстрації голосового сигналу дозволяє підвищити якість ознак шаблону користувача, що формуються традиційно.

Використання фазових даних для оцінки елементів шаблону для системи голосової аутентифікації — це додаткова інформація для підвищення якості систем голосової аутентифікації користувача.

Наступні процедури пов'язані з використанням співвідношень розрахунку основних елементів шаблону користувача системи голосової аутентифікації, а саме, частоти основного тону, формантної інформації та MFCC по амплітудно-частотній та фазовій інформації голосового сигналу.

Заключний етап обробки на фіксованій частоті дискретизації — порівняння результатів, отриманих в процесі аналізу амплітудно-частотної та фазової інформації. У деяких випадках для порівняння використовувався нормований коефіцієнт кореляції. При цьому оцінки отримані за амплітудно-частотними даними приймалися як еталон.

Після цього виконувалося проріджування аналізованих даних в два рази, наприклад, до 32 кГц, і цикл обробки повторювався. Зауважимо, що такий підхід дозволяє оцінити потенційний вплив частоти дискретизації голосового сигналу на якість формованих ознак шаблону користувача.

Обробка експериментального голосового сигналу виконувалася в системі комп'ютерної математики MatLab.

На закінчення представляються результати модельного експерименту за оцінками частоти основного тону, формантної інформації та MFCC за амплітудно-частотною та фазовою інформацією голосового сигналу.

Література:

1. Головин А.В., Исаев А.А., Мазуров В.А., Поляков В.В., Сидоренко Т.В. Уголовно-правовые и криминологические проблемы защиты информации. Алматы: Изд.центр ОФППИ Интерлигал. 2008. — 338 с.
2. ГОСТ Р-ИСО/МЭК 17799-2005. Информационная технология. Практические правила управления информационной безопасностью. М.: Стандартиформ. 2006.
3. Иконин С.Ю., Сарана Д.В. Система автоматического распознавания речи SPIRIT ASR Engine // Цифровая обработка сигналов. 2003. №4. С. 5–13.
4. Mariethoz J., Herve B., BenZeghiba M.F. Speaker Verification Based on User-Customized Password // IDIAP Research Report 01-13. Martigny. 2001. 22 p.
5. Pastushenko M., Faizulaieva O. Employment of phase characteristics of user voice signal in authentication systems. 2016 Third International Scientific-Practical Conference Problems of Infocommunications Science and Technology (PIC S&T). Kharkiv. 2016. pp. 205-206.
6. Pastushenko M., Pastushenko V. and Pastushenko O. Specifics of Receiving and Processing Phase Information in Voice Authentication Systems. 2019 International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T). Kyiv. Ukraine. 2019. pp. 621-624.